

Fluency Heuristic: A Model of How the Mind Exploits a By-Product of Information Retrieval

Ralph Hertwig and Stefan M. Herzog
University of Basel, Switzerland

Lael J. Schooler
Max Planck Institute for Human Development, Berlin

Torsten Reimer
University of Maryland, College Park

Boundedly rational heuristics for inference can be surprisingly accurate and frugal for several reasons. They can exploit environmental structures, co-opt complex capacities, and elude effortful search by exploiting information that automatically arrives on the mental stage. The fluency heuristic is a prime example of a heuristic that makes the most of an automatic by-product of retrieval from memory, namely, retrieval fluency. In 4 experiments, the authors show that retrieval fluency can be a proxy for real-world quantities, that people can discriminate between two objects' retrieval fluencies, and that people's inferences are in line with the fluency heuristic (in particular fast inferences) and with experimentally manipulated fluency. The authors conclude that the fluency heuristic may be one tool in the mind's repertoire of strategies that artfully probes memory for encapsulated frequency information that can veridically reflect statistical regularities in the world.

Keywords: fluency, fluency heuristic, recognition heuristic, ecological rationality, ACT-R

Supplemental Materials: <http://dx.doi.org/10.1037/a0013025.supp>

The human mind has long been regarded as tailored to register and to exploit frequencies of occurrence. Take, for instance, David Hume's (1740/1978) view, expressed in *A Treatise of Human Nature*:

As the habit, which produces the association, arises from the frequent conjunction of objects, it must arrive at its perfection by degrees, and must acquire new force from each instance, that falls under our observation. The first instance has little or no force: The second makes some addition to it: The third becomes still more sensible; and 'tis by these slow steps, that our judgment arrives at a full assurance. (p. 130)

Hume (1740/1978) believed that the mind unconsciously tallies frequencies and apportions degrees of belief (for Hume, the *vivacity* of an idea). He held that the mechanism for converting observed frequency into belief was finely tuned: "When the chances or experiments on one side amount to ten thousand, and on the other

to ten thousand and one, the judgment gives the preference to the latter, upon account of that superiority" (p. 141). Today, we know that Hume endowed the human mind with too exquisite a sensitivity to frequencies and that there is typically no direct mapping of environmental frequencies onto degrees of beliefs. Yet, numerous treatments of human cognition and memory center on the effects of environmental frequencies and repetition on memory and judgment (e.g., Dougherty, Gettys, & Ogden, 1999). Various mental tools exploit the "force" of these repetitions on memory. One of these mental tools is the *ACT-R fluency heuristic* that cashes in on *retrieval fluency* (Schooler & Hertwig, 2005). In what follows, we describe the fluency heuristic, and its close relative the recognition heuristic, before discussing their intellectual antecedents.

The Fluency Heuristic and the Recognition Heuristic

The research program on fast and frugal heuristics (Gigerenzer, Todd, & the ABC Research Group, 1999) has demonstrated that a user of heuristics who invested modest amounts of cognitive effort—for instance, in terms of searching for information and integrating it—could nevertheless achieve high levels of performance. Arguably the simplest of these heuristics, the *recognition heuristic* is a key example of just how far a little cognitive effort can go. It predicts which of two objects, *a* or *b*, has the higher value on a quantitative criterion. The heuristic's policy states:

If one of two objects, *a* or *b*, is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion.

Ralph Hertwig and Stefan M. Herzog, Department of Psychology, University of Basel, Switzerland; Lael J. Schooler, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany; Torsten Reimer, Department of Communication, University of Maryland, College Park.

Ralph Hertwig was supported by Swiss National Science Foundation Grants 100013-107741/1 and 100014-118283/1. Our thanks go to Julian Marewski, Thorsten Pachur, and Tim Pleskac for many constructive comments. We also thank Laura Wiles for editing the manuscript, and Gregor Caregnato, Arne Fesche, and Renato Frey for conducting the experiments.

Correspondence concerning this article should be addressed to Ralph Hertwig, Department of Psychology, Missionsstrasse 60/62, CH-4055 Basel, Switzerland. E-mail: ralph.hertwig@unibas.ch

Recognition knowledge (i.e., knowledge of the previously experienced or believed to be previously experienced) is cognitively inexpensive. In the process of retrieving a memory record, it arrives automatically and instantaneously on the mental stage, thus ready to enter inferential processes when other knowledge still awaits retrieval (Pachur & Hertwig, 2006). Low-cost recognition information, of course, will not inevitably result in high levels of accuracy. For good performance, recognition needs to be correlated with the criterion to be inferred, an issue to which we return shortly. Since Goldstein and Gigerenzer (2002) proposed the recognition heuristic, numerous studies have demonstrated that recognition is an important piece of information across various inferential tasks such as the prediction of outcomes at sport events (e.g., Pachur & Biele, 2007; Serwe & Frings, 2006) and the judgment of demographic, geographic, and biological quantities (e.g., Pohl, 2006; Reimer & Katsikopoulos, 2004; Richter & Späth, 2006).¹

Heuristics are not all-purpose inferential tools. Rather, they are applicable under limited circumstance that, ideally, can be defined. This is the case for the recognition heuristic: It cannot be applied when both objects are either recognized or unrecognized. If both objects are recognized, one applicable strategy is the fluency heuristic (Schooler & Hertwig, 2005; see Gigerenzer & Goldstein, 1996, for other heuristics) that exploits ease of retrieval and can be expressed as follows:

If two objects, *a* and *b*, are recognized, and one of two objects is more fluently retrieved, then infer that this object has the higher value with respect to the criterion.

Like the recognition heuristic, the fluency heuristic is useful whenever there is a substantial correlation—in either direction—between a criterion and recognition and/or retrieval fluency. For simplicity, we assume that the correlation is positive. The fluency heuristic relies on one inexpensive piece of mnemonic information to make an inference, namely, the fluency with which memory records (of the objects' names) are retrieved from long-term memory. That is, even if two objects are recognized, the fluency with which the names are retrieved may be different. Such differences can be exploited to make inferences about other properties of the objects.

Schooler and Hertwig (2005) implemented the fluency heuristic and the recognition heuristic within the ACT-R cognitive architecture (Anderson et al., 2004; Anderson & Lebiere, 1998), thereby being able to precisely define fluency in terms of the time it takes to retrieve memories (or chunks, to use the ACT-R terminology). ACT-R makes the assumption that information in long-term memory is stored in discrete chunks and that retrieval entails search through these to find the one that achieves some processing goal of the system. The explanatory power of the approach depends on the system's estimates of the probability that each record in long-term memory is the one sought. These probabilities are encoded in a record's *activation*. Activation tracks environmental regularities, such as an object's frequency and recency of occurrence (and is also a function of parameters such as decay of activation over time). Therefore, activation differences partly reflect frequency differences, which, in turn, may be correlated with differences in objective properties of objects. A cognitive system may be able to capitalize on differences in activation associated with various objects by gauging how it responds to them. Two phenomenological responses that are correlated with activation in

ACT-R are (a) whether a record associated with a specific object can be retrieved and (b) how quickly the record can be retrieved (henceforth, *retrieval fluency*). The first binary response was the basis for Schooler and Hertwig's implementation of the recognition heuristic. The second continuous response provided the basis for their implementation of the fluency heuristic and the existence proof that a cognitive system that relies on fluency can make moderately accurate inferences about real-world quantities. It does so by tapping indirectly—via retrieval fluency—into the environmental frequency information locked in the chunks' activation values.

Here are the goals of the current article: Schooler and Hertwig's (2005) ACT-R analysis of the fluency heuristic was theoretical in nature. From their analysis, however, follow four empirical questions that we aim to answer: First, does retrieval fluency correlate at all with objective properties of the world (Study 1)? Second, how accurately can people discriminate between often-minute differences in retrieval fluency (Study 2)? Third, to what extent do individuals' inferences actually agree with the fluency heuristic (Study 3)? Fourth, is there direct experimental evidence that fluency guides inferences about real-world quantities (Study 4)? Before we turn to these studies, we review some intellectual roots of the fluency heuristic.

Fluency, the Fluency Heuristic, and the Availability Heuristic

There are many different variants of fluency, including processing fluency and the distinctions between absolute and relative fluency, conceptual and perceptual fluency (see Alter & Oppenheimer, 2007; Reber, Schwarz, & Winkielman, 2004; Winkielman, Schwarz, Fazendeiro, & Reber, 2003). The fluency heuristic as investigated here focuses on retrieval fluency, a proximal cue that can inform and influence human inference across a wide range of target criteria. It has been demonstrated to underlie, for instance, a person's memory of the past and prediction of future memory recall performance (Benjamin, Bjork, & Hirshman, 1998; Benjamin, Bjork, & Schwartz, 1998); an eyewitness's confidence in her or his memory (Shaw, 1996; Shaw, McClure, & Wilkens, 2001); assessments of one's ability to learn (e.g., Koriat & Ma'ayan, 2005); and people's confidence in their general knowledge (e.g., Kelley & Lindsay, 1993; Unkelbach, 2007). It has also been invoked in explaining consumer decisions (e.g., Schwarz, 2004).

Schooler and Hertwig's (2005) use of the term *fluency heuristic* harkens back to a long research tradition on fluency, in particular the work by Jacoby and Dallas (1981); Kelley and Jacoby (1998); Kelley and Lindsay (1993); Whittlesea (1993); and Whittlesea and Leboe (2003). Abstracting from the different meanings of the term *fluency heuristic* across these articles, the gist involves three properties: (a) the attribution of "fluent" processing to prior experience, (b) the resulting conscious experience of familiarity, and (c) the

¹ At the same time, the assumption that no other probabilistic information beyond recognition will be used if the recognition heuristic is applicable has been vigorously challenged (e.g., Bröder & Eichler, 2006; Newell & Fernandez, 2006; Oppenheimer, 2003; Pohl, 2006; Richter & Späth, 2006; but see also Pachur, Bröder, & Marewski, 2008; Pachur & Hertwig, 2006).

assumption that relative fluency can be used as a basis for recognition memory.²

A second intellectual root of the fluency heuristic, as studied here, is the availability heuristic, one of the key heuristics proposed and investigated in the heuristics-and-biases research program (Tversky & Kahneman, 1974). Two interpretations of this heuristic have emerged, one of which includes the notion of retrieval fluency (Tversky & Kahneman, 1973, pp. 208, 210; see also Schwarz et al., 1991; Schwarz & Wänke, 2002). In one version, the availability heuristic rests on the *actual frequencies* of instances or occurrences retrieved, for instance, the occurrences of heart attack among one's acquaintances to assess the risk of heart attack among middle-aged people. Another rests on the ease, that is, *fluency* with which the operation of retrieval of these instances and occurrences can be performed (for more on the distinction between these two notions of availability, see Hertwig, Pachur, & Kurzenhäuser, 2005; Sedlmeier, Hertwig, & Gigerenzer, 1998).

Although there are similarities between the fluency heuristic and the latter version of the availability heuristic, fluency researchers have conceptualized them as two distinct heuristics (Jacoby & Dallas, 1981, p. 701; Jacoby & Whitehouse 1989, p. 127). One difference concerns the entities that are being retrieved. The availability heuristic derives its assessment of the frequency (probability) of the target event, say, risk of heart attack among middle-aged people, from the effortlessness (or lack thereof) with which prior instances of the target event could be retrieved (Tversky & Kahneman, 1974). The fluency heuristic, as defined by Schooler and Hertwig (2005), by contrast, bases its inferences simply on the speed with which the event category itself (e.g., myocardial infarction) is recognized. We realize, however, that the extent to which this property renders the fluency and the availability heuristics distinct depends on one's definition of availability. If availability extends to the retrieval of the event category itself, the fluency heuristic and the availability heuristic will be indistinguishable. Let us therefore reiterate the view we expressed in Schooler and Hertwig (2005, p. 626):

We would have no objection to the idea that the fluency heuristic falls under the broad rubric of availability. In fact, we believe that our implementation of the fluency heuristic offers a definition of availability that interprets the heuristic as an ecologically rational strategy by rooting fluency in the informational structure of the environment.

Moreover, we believe that researchers in both the availability heuristic and the fluency heuristic traditions will find the present set of studies relevant. If one interprets the heuristics as distinct, our studies will foster the understanding of the predictive power of the fluency heuristic. Researchers interpreting availability and fluency as two sides of the same coin may find the present studies enriching to the extent that they provide one precise definition of availability and investigate its predictive power in making inferences about the world. That predictive power or lack thereof is, in fact, the focus of Study 1.

Study 1: Is It Worth Exploiting Retrieval Fluency?

How can one learn the association between retrieval fluency and a criterion when the criterion is not accessible? In the context of the recognition heuristic, Goldstein and Gigerenzer (2002) proposed that there are "mediators" in the environment that both

reflect (but do not reveal) the criterion and are accessible to decision makers' senses. For example, one may have no direct information about the ability of a tennis player, say, Roger Federer. Yet, his strength as a player may be reflected by how often his name is mentioned in the newspaper (*ecological correlation*). Because the newspaper is accessible, it can operate as a mediator. The frequency of mentions in the newspaper, in turn, is correlated with how likely someone is to recognize the player's name (*surrogate correlation*). Finally, how good or poor a proxy a person's recognition knowledge is of the criterion is captured in the *recognition validity*. Based on this chain of correlations between the criterion, the mediator, and the mind, a person would be able to make inferences about a player's strength depending on whether he or she recognized his name.

The same ecological analysis can also be conducted for the fluency heuristic, except that retrieval fluency replaces recognition. Frequency of mentions in the newspaper may be correlated not only with recognition but also with how quickly a person can retrieve the memory record representing the name Roger Federer. That is, the mediator can influence the recognition latency (thus giving rise to varying degrees of fluency) and the probability of recognition. Fluency, however, can be high for the wrong reasons and need not reflect environmentally valid variations in exposure. It may be sensitive to recent exposures, thus compromising the link between fluency and the criterion.

So, what is the ecological validity of retrieval fluency? Investigations of the chain of correlations between the criterion, frequency of mentions, and retrieval fluency in a clearly defined reference class of objects do not exist.³ One rare exception is the investigation by Alter and Oppenheimer (2006). They found that the complexity of a share's name and the pronounceability of companies' three-letter stock ticker codes—both variables are indicators of the verbal fluency of a share's name—are predictive of the actual performance of those shares in the stock market immediately after their release onto the stock exchange. This analysis is particularly interesting insofar as fluency in their investigation is not just an indicator of the criterion, the share's performance on the stock market, but appears to causally determine stock performance in the short term.

² Whittlesea and Leboe (2003; Whittlesea, 1993) make an important distinction between two kinds of fluency heuristics. In their view, a person's feeling of fluency can either be a reflection of enhanced processing, or, alternatively, a person's perception of fluency can be relative to his or her "fluency" expectations. We focus on the first notion of fluency. Moreover, we also do not address the role of naïve theories that may trigger the use or disuse of the fluency heuristic (see Schwarz, 2004).

³ Two experts on fluency, P. Winkielman (personal communication, October 13, 2007) and C. Unkelbach (personal communication, October 12, 2007) confirmed our impression that there are hardly any systematic analyses of the ecological validity of fluency in general and retrieval fluency in particular. Ecological analyses remain rare, however, even if one equates availability and fluency. Although Tversky and Kahneman (1973) stressed that "availability is an ecologically valid clue for the judgment of frequency because, in general, frequent events are easier to recall or imagine than infrequent ones" (p. 209), many subsequent studies have focused on circumstances under which reliance on availability leads astray (for a systematic analysis of the validity of availability in a classic task, see Sedlmeier et al., 1998).

In Study 1, we investigated whether retrieval fluency, like recognition, is a proxy for real-world quantities across five different reference classes in which we expected retrieval fluency to be effective. In this and the subsequent studies, we could not directly measure retrieval fluency, so instead we collected recognition latencies, that is, how long it took participants to judge an object as recognized or not. Of course, recognition latency is not a perfect proxy for retrieval fluency. Recognition latency includes other systematic and unsystematic components such as the time it takes to read a word, to decide whether it is recognized, and to output a motor response. It may well be that these factors drown out the contribution that retrieval latency makes to the overall recognition time.

Method

Participants. One hundred and sixty students from the University of Basel participated in the study (98 women and 62 men, mean age = 24.9 years), which was conducted at the Department of Psychology. Participants received either money (7.50 Swiss Francs = US\$6.10) or a course credit for their participation.

Material. To study the predictive power of fluency across different domains, we compiled five different environments (see supplemental materials): (a) the cities environment comprising all 118 U.S. cities with more than 100,000 inhabitants (Butler, 2003); (b) the companies environment containing all 100 German companies with the highest revenue in 2003 (“Die 100 grössten deutschen Unternehmen,” n.d.); (c) the music artists environment including all 106 most successful artists in the U.S., in terms of the cumulative sales of recordings in the U.S. from 1958 to 2003 (“Top artists,” 2003); (d) the athletes environment including the 50 richest athletes in 2004 (“The Best-Paid Athletes”, 2004); and (e) the billionaires environment including the 100 wealthiest people in 2004 (“The World’s Richest People”, 2004). For each environment, we first determined the frequencies of mentions in the media (the mediator), by using COSMAS, the largest online archive of German print media (e.g., encyclopedias, books, and daily and weekly newspaper articles).⁴

Procedure. We recorded people’s recognition latencies for each of the total of 474 objects. To avoid exhaustion, each participant saw a subset of the 474 objects. Forty participants were presented with the names of 100 German companies, 106 music artists, 50 sportspeople, and 100 billionaires, one at a time on the computer screen.⁵ The order of objects within each environment was randomized across participants, as was the order of environments. A second group of 120 participants saw approximately a third of a pool of 525 cities drawn from different regions of the world, including 118 U.S. cities with more than 100,000 inhabitants (see the material used in Volz et al., 2006). Because the set of U.S. cities is the largest and the only one that encompasses all objects within a defined size range, we focus on the U.S. set in the following analysis. The names of the objects were presented one at a time (in random order), and respondents were asked to decide whether they had heard of the object. Both groups of participants indicated their decisions by pressing one of two keys, with the assignment of keys to responses counterbalanced across participants. In addition, respondents were instructed to make their decisions as quickly and as accurately as possible. The time that elapsed between the presentation of the object’s name and the participant’s keystroke was measured. After each response, the

screen remained blank for 1,000 ms. Participants responded to 10 practice trials.

Results

How ecologically valid is retrieval fluency? To quantify the strength of the relationship between the environmental criteria, the frequency with which the names of the objects were mentioned in the media, and retrieval fluency (measured in terms of recognition latency), we calculated three measures: the fluency validity, the ecological validity, and the surrogate validity. The strength of the relationship between retrieval fluency and the criterion is defined as the proportion of times a faster recognized object has a higher criterion value than an object requiring more time to be recognized (in a given reference class). The fluency validity v_f is thus:

$$v_f = R_f / (R_f + W_f),$$

where R_f is the number of correct (right) inferences made by the fluency heuristic computed across all pairs in which the difference in recognition latencies between two recognized objects equals or exceeds a just noticeable difference (JND), and W_f is the number of incorrect (wrong) inferences under the same circumstances. Note that validity is linearly related to Goodman-Kruskal’s γ , an ordinal measure of association. Concerning the JND, Schooler and Hertwig (2005) assumed that there are limits to people’s ability to discriminate between retrieval times. They assumed that the difference needs to be as large or larger than a JND of 100 ms, otherwise people resort to guessing. This value of 100 ms was based on Fraisse’s (1984) extensive review of the timing literature—on his conclusion that durations of less than 100 ms are perceived as instantaneous—and is consistent with our findings in Study 2. Although all analyses reported in this article assume a JND of 100 ms, results do not depend on this specific value.

Ecological validity describes the relation between the criterion to be inferred and how often an object was mentioned in the media (its environmental frequency, measured with COSMAS). Ecological validity is defined as the proportion of times an object with the higher number of mentions in the media has a higher criterion value than does an object with a lower number of mentions. The ecological validity v_e is thus:

$$v_e = R_e / (R_e + W_e),$$

where R_e is the number of correct cases and W_e the number of incorrect cases computed across all pairs where one object occurs more frequently in the media than does the other.

⁴ We used print material published between 2000 and 2004. COSMAS can be accessed on the Internet: <http://corpora.ids-mannheim.de/~cosmas/>

⁵ For the analyses of Studies 1–3, we excluded the names of eight companies and four musicians because of their lengths (e.g., “Bob Seger and the Silver Bullet Band”), that is, the risk of measuring reading times rather than recognition latencies. We also checked whether there is a correlation between the lengths of names (i.e., number of letters) and the criterion in question; we found none. For 23 of the 92 German companies, the official company name is different from the colloquial name. For these companies, the media mentions were substantially higher for the colloquial names than for the official names. Therefore, the ecological analyses rest on the former.

Finally, the surrogate validity describes the relation between the mediator and recognition latencies. It is defined as the proportion of times an object with a higher number of mentions in the media is retrieved faster than an object with a lower number of mentions. The *surrogate validity* v_s is thus:

$$v_s = R_s / (R_s + W_s),$$

where R_s is the number of correct cases and W_s the number of incorrect cases computed across all pairs where one object occurs more frequently in the media than does the other (this calculation also assumes a JND of 100 ms). Goldstein and Gigerenzer (2002) and Schooler and Hertwig (2005) described the relations between criterion, mediator, and mind in terms of one validity measure and two correlation measures (ecological correlation and surrogate correlation). To simplify comparisons between these correlations, we consistently quantify them in terms of one currency, validity. All three validities can be interpreted as the conditional probability that object a scores higher (or lower) on one dimension than does object b , given that object a scores higher (or lower) on a second dimension than does object b .

Table 1 shows, separately for each environment, the median response time, the average and median number of occurrences in COSMAS per object, and the recognition rates. Across environments, the capacity to recognize objects varied widely, ranging from 68% (music artists) to merely 11% (billionaires) recognized objects (Table 1). Mirroring their low recognition rates, the median environmental frequencies of the names of the billionaires and athletes, respectively, is a magnitude smaller than that in the other domains. Finally, except in the billionaires environment, the median time for recognized objects was markedly shorter than the response time for unrecognized objects.

To what extent is retrieval fluency a proxy for inferring real-world quantities? Figure 1 depicts the relations between the criterion, the frequency with which the names of the objects occurred in the media, and recognition latency and recognition, respectively. In all five environments, fluency validity exceeded chance level (.50), ranging from .66 in the cities environment to .58 in the music artists and companies environments, respectively. Figure 1 also shows that recognition validity consistently exceeds fluency validity—an observation to which we return in the final discussion.

Figure 2 plots fluency validity as a function of the magnitude of the objective differences in latencies (summarized for four bins: 0–99 ms, 100–399 ms, 400–699 ms, >700 ms), separately for the

five environments. There is a clear tendency that the larger the objective difference, the higher the validity of fluency. This tendency follows from the ACT-R framework, in which the activation of a memory record tracks environmental frequencies. Objects with larger criterion values tend to occur more frequently in the environment, and thus their memory records tend to be more quickly retrieved. Consequently, large differences in latencies are likely to represent a pair of objects in which one object has a large criterion value and one has a small value. For such pairs, fluency can be expected to be quite valid.

To conclude, we found that differences in recognition latencies are indicative of criteria across five different environments. The strength of the relationship varies across environments. In the cities environment, for which we recorded the strongest relationship between the mediator and the criterion, we also observed the highest fluency validity. Similarly, environments with low ecological validity such as the companies and the music artists environments also yielded relatively low levels of fluency validity. Based on this ecological analysis of fluency across five environments, we can now conclude that, based on retrieval fluency, one can at least theoretically infer distal properties of the world.

Study 2: Are People Able to Exploit Retrieval Fluency?

To exploit retrieval fluency, people need to be able to judge accurately whether recognizing object a 's name takes longer than recognizing object b 's name, or vice versa. For instance, can a person tell the difference between, say, instantaneously recognizing *BMW*, and taking a moment to recognize *Siemens*? Such differences in recognition latencies partly reflect retrieval time differences; these in turn reflect the base-level activation of the corresponding memory records, which tracks environmental frequency and recency (Schooler & Hertwig, 2005). In Study 2, we investigated the extent to which people can accurately tell such differences apart. In this and Study 3, we excluded the athletes and the billionaires environments, because the low recognition rates (26% and 11%, see Table 1) make it difficult to investigate the critical cases in which people recognize both objects.

Method

Participants. One hundred and twenty students (81 women and 39 men, mean age = 23.2 years) participated in the study, 60

Table 1
Response Time, Recognition Rates, and Environmental Frequencies Obtained Across the Five Environments Investigated in Study 1

Environment	Median response time per object (ms)				Recognition rate		Environmental frequencies ^a	
	Recognized		Unrecognized		<i>M</i>	<i>Mdn</i>	<i>M</i>	<i>Mdn</i>
	<i>M</i>	<i>Mdn</i>	<i>M</i>	<i>Mdn</i>				
Cities	887	835	1129	1013	.53	.51	741	85
Companies	890	871	1045	999	.47	.48	671	142
Music artists	737	699	1045	1043	.68	.71	240	55
Athletes	948	920	1027	963	.26	.22	218	7
Billionaires	1591	1460	1176	1040	.11	.09	22	1

^a Measured in terms of hits in COSMAS (see Footnotes 4 and 5).

from the University of Basel and 60 from the Free University (Berlin). Participants received either money (7.50 Swiss Francs = US\$6.10 in Basel; or €10 = US\$12.50 in Berlin) or course credit for their participation.

Material. Participants responded to a total of 156 pairs of U.S. cities ($n = 59$), German companies ($n = 46$), and music artists

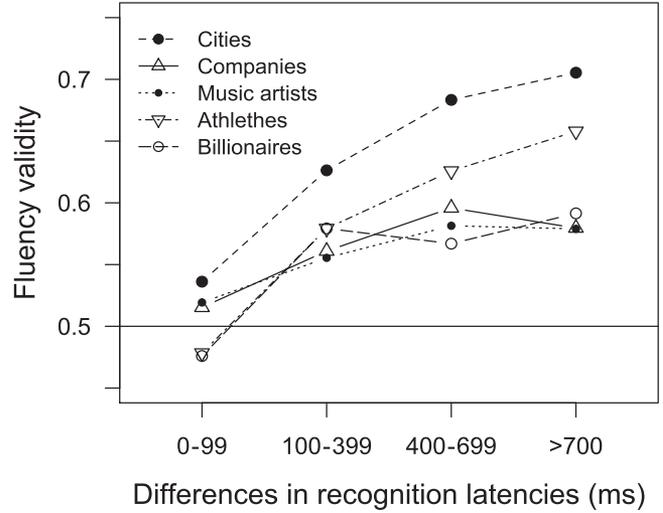
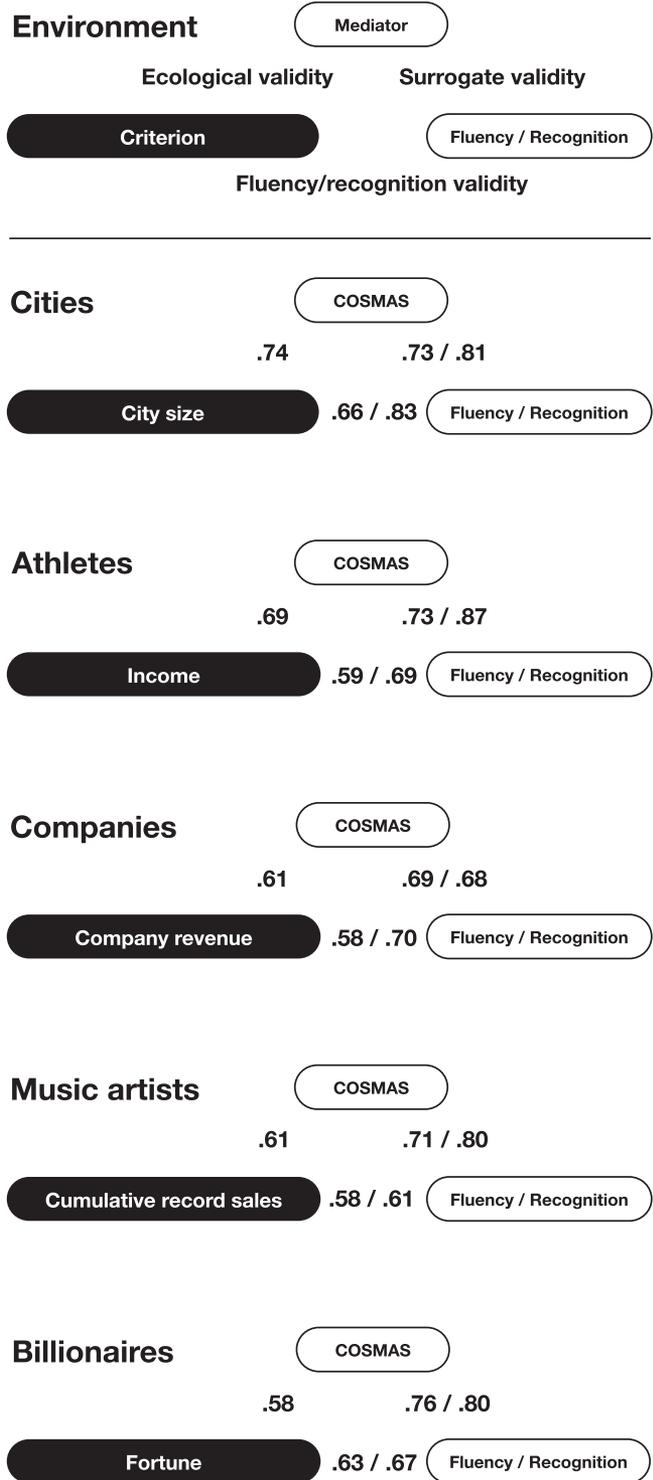


Figure 2. Increasing fluency validity accuracy as a function of increasing differences in recognition latencies. Differences in objects *a* and *b*'s recognition latencies were classified into four bins. Then, we calculated the fluency validity per bin. We chose to create bins such that each bin included judgments of nearly all participants. Other ways of creating the bins, however, result in the same general trend. The horizontal line at 0.5 represents chance level validity.

($n = 51$). Pairs were formed by randomly drawing objects (without replacement) from the sets of cities, companies, and music artists, respectively. We repeated the process of randomly pairing objects (separately for each environment) a total of 30 times. Each of the 30 random pairings was then presented to 4 participants.

Procedure. Participants were informed that they would be presented with 156 pairs of objects. Pairs were blocked for each environment, and order of environments was counterbalanced. Participants saw the name of the first object on the computer screen. By pressing one of two buttons, they indicated whether they recognized it. Then they saw the second name and responded accordingly. The order of the names within each pair was randomized. Immediately following the recognition task, the second task probed participants' ability to discriminate between the response times for the two objects in question. Both names were presented on the screen, next to each other (the location, left vs. right, was randomly determined). Respondents indicated for which of the two names they more quickly arrived at the previous judgment

Figure 1. Ecological analysis of fluency validity and recognition validity across five environments. The triangles, adapted from Goldstein and Gigerenzer (2002; Figure 7), show the relationship between criterion, mediator, and the mind, measured in terms of validity. The inaccessible criterion is reflected but not revealed by the mediator variable (ecological validity). The mediator (environmental frequencies as measured by frequency of occurrences in COSMAS) influences retrieval speed (and probability of recognition); this link is expressed in terms of the surrogate validity. The mind in turn may use recognition speed or recognition to infer the criterion (the accuracy of this inference is captured in terms of fluency or recognition validity). Note that the ecological validity (the relationship between mediator and criterion) is independent of whether the mind uses retrieval speed or recognition.

(regardless of whether the judgment was “recognized” or “unrecognized”). For half of the participants, the order of recognition and discrimination task was reversed. That is, they were presented with the two names and asked for which they arrived at a covert recognized versus unrecognized judgment more quickly (covert insofar as they were not asked for an overt recognition judgment). Once this discrimination was made, objects were presented again, one at a time, and respondents indicated whether—a moment ago—they had recognized the name. Participants first responded to six practice trials.

Results

First, we briefly describe the recognition judgments. When recognition preceded discrimination, the rates of recognition and the response times replicated those obtained in Study 1 (compare Table 1 and Table F in the supplemental materials). As in Study 1, recognized responses are, on average, markedly faster than unrecognized responses. When recognition succeeded discrimination, a similar picture emerged, with one exception. Response times, both for recognized and unrecognized objects, were much shorter than those when recognition preceded discrimination; recognition rates, however, were almost identical for both orders.

To measure how accurately people can discriminate recognition latencies, we compared participants’ *objective recognition latencies* (i.e., the time it took them to judge an object to be recognized) with their judgments of which of the two objects was recognized faster. For this analysis, we focused on the condition in which recognition came first and discrimination second, and recognition latencies were not biased through the objects’ previous presentation. Table 2 reports how accurately participants were able to discriminate differences in latencies longer than 100 ms. Across all environments, participants scored a median .76 accuracy, with quite comparable levels for all three environments (.78, .77, and .74, respectively). In addition, Figure 3 plots the *discrimination accuracy* (i.e., proportion of correct discriminations) as a function of the magnitude of the objective differences in latencies. The larger the objective difference, the better participants are able to discriminate. When differences are shorter than 100 ms, accuracy drops close to chance level ($M = .53$; $Mdn = .54$; 95% confidence interval (CI) = .49, .58), consistent with Fraisse’s (1984) con-

clusion that latencies that differ by less than 100 ms are perceived as instantaneous. When differences exceeded 700 ms, the average discrimination accuracy rose to .86 and higher. In light of this ability to discriminate, it is not surprising that if one calculates fluency validity, using participants’ subjective judgments of which of the two objects was recognized faster, the subjective fluency validities—.65, .58, and .52 for the cities, companies, and music artists environments, respectively—track closely the fluency validities derived by using objective differences (Table 2). Note that in the music artists environment the confidence interval includes chance performance of .50.

The results in Figure 3 also suggest that participants’ ability to discriminate is particularly pronounced when they stand to benefit most from relying on fluency, namely, when the differences in recognition latency are large (see Figure 2). Indeed, as Table 3 shows, large differences in recognition latency give rise to high levels of discrimination accuracy and fluency validity in the cities environment. In the other two environments, the results exhibit the same trend, but the trend is less pronounced.

In sum, we observed three results. First, people prove to be quite good at discriminating between recognition latencies whose difference exceeds 100 ms. Second, even when taking less-than-perfectly accurate discriminations into account, subjective fluency judgments are a moderately good predictor of the criterion, except in environments in which ecological validity of fluency information is low to begin with (e.g., music artists environment). Last, we found that people’s ability to discriminate is highest for those pairs in which the validity of fluency peaks, as suggested by the ACT-R framework.

Study 3: Are People’s Inferences in Line With the Fluency Heuristic?

The next study addresses three issues: First, we investigated the extent to which people’s inferences are in line with the fluency heuristic and the recognition heuristic. Second, we examined the extent to which differences in retrieval fluency affect people’s accordance to the fluency heuristic. Third, we tested the prediction that inferences agreeing with the fluency heuristic require less time, relative to inferences conflicting with the fluency heuristic. Let us briefly develop the latter two issues.

Table 2
Applicability and Validity of the Fluency and Recognition Heuristics, Respectively, Across Three Environments

Environment	Fluency heuristic								Recognition heuristic			
	Fluency applicability		Discrimination accuracy ^a		Objective fluency validity ^a		Subjective fluency validity		Recognition applicability		Recognition validity	
	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI
Cities	.20	.18, .23	.78	.75, .82	.64	.59, .68	.65	.62, .68	.48	.47, .50	.82	.80, .83
Companies	.22	.20, .25	.77	.72, .81	.58	.53, .64	.58	.55, .62	.47	.45, .49	.69	.67, .71
Music artists	.31	.29, .33	.74	.70, .78	.55	.51, .58	.52	.50, .55	.40	.38, .42	.56	.54, .58

Note. Discrimination accuracy reflects how accurately respondents can discriminate between the recognition latencies of two objects. Fluency validity is simulated using both the objective recognition latencies (objective fluency validity) and people’s judgments (subjective fluency validity) of which object was recognized faster. CI = 95% confidence interval.

^a Only calculated from the participants in the condition in which recognition precedes discrimination.

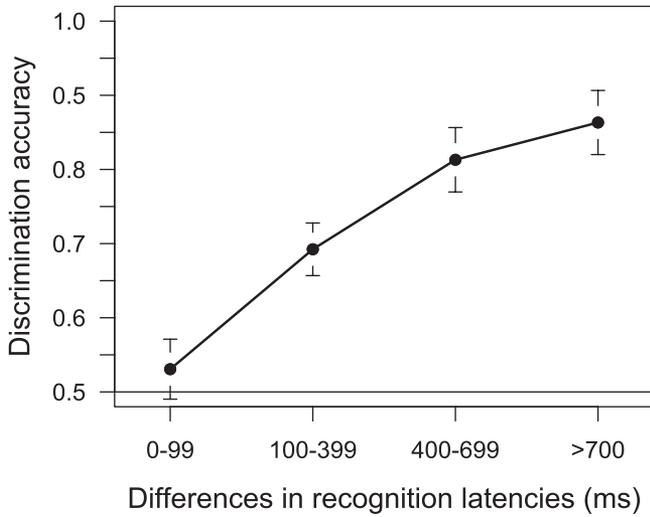


Figure 3. Increasing discrimination accuracy as a function of increasing differences in recognition latencies. Differences in objects *a* and *b*'s recognition latencies were classified into four bins. Then, we calculated the mean discrimination accuracy per bin (across participants). Error bars indicate 95% confidence intervals. The horizontal line at 0.5 represents chance level accuracy.

Does Accordance to the Fluency Heuristic Increase as a Function of the Difference in Recognition Latencies?

In Studies 1 and 2, we found that the larger the differences between the two objects' recognition latencies, the more likely that the exploitation of fluency leads to a correct inference (Figure 2) and the more accurate the discrimination between them (Figure 3). Rather than assuming that differences beyond 100 ms give rise to the same degree of fluency heuristic accordance, adherence may depend on the magnitude of the difference.

Do Inferences That Agree With the Fluency Heuristic Require Less Time?

Pachur and Hertwig (2006) predicted and found evidence that inferences consistent with the recognition heuristic require less time than do inconsistent inferences. This prediction can also be

extended to fluency information. Recognizing an object requires the memory record of the object to be retrieved, and the speed with which this process unfolds can be used as a proxy for people's senses of fluency. Therefore, fluency information is essentially produced at first sight of the names of the objects. It is available and ready to enter the inferential process while other information has yet to be retrieved. In contrast, inferences inconsistent with the fluency heuristic need to rely on information beyond recognition and recognition speed. Unless the result of mere guessing, such inferences rely on the retrieval of probabilistic cues or explicit knowledge about the criterion. This logic suggests that inferences based on fluency information may be made faster than inferences inconsistent with the fluency heuristic. Yet, if reliance on fluency requires additional checks such as whether the difference in latencies is "good enough," then fluency-based inference may turn out not to be made faster.

Method

Participants and design. Eighty students from the Free University (Berlin) participated in the study (42 women and 37 men; 1 participant failed to indicate gender information; mean age = 25.3 years). As in Study 2, participants were presented with 156 pairs of U.S. cities (*n* = 59), German companies (*n* = 46), and music artists (*n* = 51), and within each pair were asked to choose the object with the higher value on the criterion (*inference task*). Furthermore, each participant indicated which objects he or she recognized (*recognition task*). Half of participants took this recognition test before the inference task and half after. They received a flat fee of €10 (US\$12.50).

Material. For the inference task, items were constructed in a similar way as in Study 2. For each environment, we randomly created 40 lists of pairs of objects (with the constraint that each object could occur only once in a list). Each of these lists was then presented to 2 participants; 1 received the recognition test before and the other after the inference task. Pairs of objects were presented in three blocks (representing the three environments). Once all inferences were made, the individual objects were again presented in three environmental blocks (in the same order as in the inference task). Participants were asked to report whether they recognized the name. The order of the blocks, the order of objects within each block, and the location of objects within each pair (left

Table 3
Mean Discrimination Accuracy and Mean Fluency Validity Across Three Environments as a Function of Small Versus Large Differences (Δ) in Recognition Latencies

Environment	Discrimination accuracy				Fluency validity			
	Small Δ	Large Δ	Difference		Small Δ	Large Δ	Difference	
			<i>M</i>	CI			<i>M</i>	CI
Cities	.72	.84	.12	.06, .18	.55	.71	.16	.08, .23
Companies	.67	.84	.17	.09, .25	.56	.60	.04	-.08, .15
Music artists	.67	.80	.13	.08, .18	.53	.57	.04	-.03, .11

Note. Across all pairs of objects and within each person, we calculated all differences in the recognition latencies for objects *a* and *b*. Using a person's median value, we classified all differences into sets of *large* and *small* ones, respectively, and calculated the average discrimination accuracies and fluency validities, averaged across participants. The mean difference columns represent the mean differences between both sets, averaged across participants. CI = 95% confidence interval.

side vs. right side of the screen) were determined at random for each participant.

Procedure. After an introductory text explaining the inference task, pairs of objects were displayed on a computer screen. Participants indicated their inferences by pressing one of two keys. They were encouraged to respond as quickly and accurately as possible. The time that elapsed between the presentation of the objects and participants' keystroke was measured. Each inference trial began with the presentation of a fixation point (a cross in the center of the screen), followed after 500 ms by the respective pair. The names appeared simultaneously (left and right of the fixation point) and remained on the screen until a response was given. After each response, the screen remained blank for 1,100 ms. Participants first responded to six practice trials. The group of participants that had completed the inference task first then immediately took the recognition task, in which they indicated whether or not they had ever heard the name of the respective object.

Results

Recognition rates and latencies mimicked those observed in Studies 1 and 2 (see Table G in the supplemental materials), except in the companies environment, in which participants recognized more objects than in the previous studies. This boost in recognition, however, is consistent with the fact that Study 3 recruited only German students (and no Swiss students).

To what degree do people's inferences agree with the fluency heuristic? For each participant, we computed the percentage of inferences that were in line with the fluency heuristic among all cases in which it could be applied (i.e., pairs in which both objects were recognized), excluding pairs with differences in recognition latencies smaller than 100 ms. Table 4 shows fluency applicability and fluency validity across environments. The values track those obtained in Study 2 (compare Tables 2 and 4). The fluency heuristic applicability ranged between about a fourth and a third of all inferences. The mean fluency heuristic accordances (i.e., proportion of inferences consistent with the fluency heuristic) were .74, .63, and .68 in the cities, the companies, and the music artists environments, respectively.⁶ As Figure 4a shows, in all environments there was substantial interindividual variation in the proportion of judgments that agreed with the heuristic. The rate of an individual's fluency heuristic accordances ranged between 1.00 and .30 across environments. Only a few participants appear to have systematically decided against fluency: Across environments, merely 12% of participants' accordances rates were below .50. Finally, Table 4 also shows participants' average score of correct inferences (in pairs in which both objects were recognized), which ranged between .58 (companies environment) and .71 (cities environment).

Next to comparing fluency accordances with a 50% baseline, one can also test whether observed fluency accordances is higher than that expected by chance: Let us assume that a person does not make use of the fluency heuristic. By mere chance one would expect that his or her inferences and those predicted by the heuristic coincide in more than 50% of the cases (assuming the person's accuracy and fluency validity are better than chance). There are two ways in which the person's and the heuristic's inferences can coincide: (a) both choose *a* (the larger object) or (b) both choose *b* (the smaller object). The probability that both

choose *a* equals the person's level of accuracy (*acc*) times the fluency validity: $acc \times v_f$. The probability that both choose *b* is the product of the complementary probabilities: $(1 - acc) \times (1 - v_f)$. Consequently, the overall probability of the person and heuristic choosing the same object by chance (henceforth, baseline accordances) equals the sum of both products. Averaged across individuals, the rates of baseline accordances were 59%, 53%, and 52% in the cities, the companies, and the music artists environments, respectively.

For each person, we then calculated the difference between his or her baseline accordances and his or her observed fluency accordances. In the cities environment, the observed accordances was 16 percentage points ($SD = 15\%$; $CI = 12\%, 19\%$; $Mdn = 16\%$; $d = 1.03$) higher than the individual-specific baseline accordances. In the companies environment, the respective difference was 11 percentage points ($SD = 13\%$; $CI = 8\%, 13\%$; $Mdn = 12\%$; $d = 0.83$); finally, in the music artists environment the difference was 15 percentage points ($SD = 11\%$; $CI = 13\%, 18\%$; $Mdn = 15\%$; $d = 1.38$). That is, when comparing the empirically found fluency accordances against a benchmark that takes the occurrence of coincidental accordances into account, the predictive power of the fluency heuristic remains sizeable.

To what degree do people's inferences agree with the recognition heuristic? Figure 4b shows individuals' accordances rates across environments. A vast majority of inferences can be captured by the recognition heuristic, when it can be applied (see Table 4 for applicability rates). The mean proportion of recognition heuristic accordances was .91, .90, and .93 in the cities, companies, and music artists environments, respectively. Averaged across environments, 27% of respondents conformed to the recognition heuristic every time. Participants' average score of correct inferences (in pairs in which one object was recognized and the other was not) ranged between .59 (music artists environment) and .80 (cities environment).

Challenging Goldstein and Gigerenzer's (2002) assumption that the recognition heuristic takes into account only whether an object is recognized or not, Newell and Fernandez (2006, p. 333) found a negative correlation between the proportion of times a city was chosen over unrecognized cities and the speed with which participants correctly categorized the name of this city ($r = -.382$). Consistent with Newell and Fernandez and across all environments, we also found that a recognized object was more often inferred to be the larger one (relative to the unrecognized object), the faster it was recognized: cities (Spearman $r = -.343$, $p = .001$), companies ($r = -.268$, $p = .01$), and music artists ($r = -.325$, $p = .001$), respectively. Convergent evidence that recognition strength correlates with following the recognition heuristic comes from an fMRI study by Volz et al. (2006). They found that decisions in accordance with the heuristic correlate with higher

⁶ The order of the recognition and inference tasks had no statistically significant effect on the accordances to the fluency and recognition heuristics in the cities, companies, and music artists environments, respectively. None of the implications of the reported analysis changed when analyzing the two task orders separately.

Table 4
Applicability, Accordance, and Validity Rates for the Recognition and Fluency Heuristics, Respectively, Across Three Environments Investigated in Study 3

Environment	Fluency heuristic								Recognition heuristic							
	Fluency applicability		Fluency accordance		Fluency validity		Participant accuracy		Recognition applicability		Recognition accordance		Recognition validity		Participant accuracy	
	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI	<i>M</i>	CI
Cities	.23	.21, .26	.74	.71, .78	.67	.64, .70	.71	.67, .74	.47	.45, .49	.91	.89, .93	.84	.83, .86	.80	.77, .83
Companies	.30	.27, .32	.63	.60, .66	.60	.57, .63	.58	.55, .61	.45	.42, .47	.90	.87, .93	.74	.72, .76	.70	.67, .73
Music artists	.33	.31, .35	.68	.65, .70	.56	.53, .59	.59	.56, .62	.39	.37, .41	.93	.91, .96	.57	.55, .60	.59	.56, .61

Note. Participant accuracy denotes the level of inferential accuracy that respondents reached. CI = 95% confidence interval.

activation in areas of the brain that have previously been associated with greater recognition confidence.

Is accordance to fluency a function of the difference in recognition latencies? Study 2 showed that the larger the differences between the objects' recognition speed, the more accurately people can discriminate between the two objects' recognition latencies (see Figure 3). Does the regularity also relate to the likelihood with which people accord to the fluency heuristic? Figure 5 plots fluency heuristic accordance as a function of differences in recognition latencies. Accordance rates, indeed, do increase with larger differences.⁷

Do inferences that agree with the fluency heuristic and the recognition heuristic require less response time? To answer this question, for each participant we calculated the median response time—separately for each of the three environments—for inferences consistent and inconsistent with both the recognition and the fluency heuristics. Figure 6 shows the mean of the differences between these two values. Across all environments, inferences conflicting with the fluency heuristic take markedly longer than do inferences consistent with it. The same regularity also holds for the recognition heuristic, replicating Pachur and Hertwig's (2006) finding. As the confidence intervals indicate, the mean of each of the six differences is significantly greater than zero. Moreover, the effect sizes (Cohen's *d*; Cohen, 1988) for the differences in response times are greater for the recognition heuristic than for the fluency heuristic across all three environments (cities: $d = 0.74$ vs. $d = 0.54$; companies: $d = 0.62$ vs. $d = 0.48$; and music artists: $d = 0.85$ vs. $d = 0.53$).

To summarize, in about two thirds to three fourths of inferences in which the fluency heuristic was applicable, people's actual choices conformed to those predicted by the heuristic. We also found that the larger the difference between recognition latencies (for two objects), the greater the likelihood that the actual inference adheres to that predicted by the fluency heuristic. Moreover, consistent with the notion of recognition and fluency's retrieval primacy, we found that inferences that agree with both heuristics take less time than do those that conflict with the heuristics (see Pachur & Hertwig, 2006). Last but not least, it is worth pointing out that the fluency heuristic was applicable in 23% to 33% of inferences, depending on the environment (Table 4). These rates appear modest. The rate of applicability, however, is a function of both the heuristic and the environment in which it is used. If we had chosen, for instance, German or Swiss cities with more than

100,000 inhabitants, rather than U.S. cities, the rate of fluency applicability would have skyrocketed (given our German and Swiss respondents). Even when applicable, however, people would likely have not relied on any inferential tool because of their direct knowledge of the criterion variable (Pachur & Hertwig, 2006), thus rendering tests of strategies such as the fluency heuristic mute.

Study 4: Are People's Inferences in Line With the Fluency Heuristic: Experimental Evidence?

The goal of the final study is to provide evidence whether or not fluency itself, rather than other factors possibly associated with fluency (e.g., the amount of knowledge about the objects, ease of retrieving this information, etc.), can guide inferences. Specifically, in Study 3, what made respondents choose the more fluently retrieved object need not have been fluency per se, but may have been, for example, the mere amount of knowledge associated with it. Reliance on fluency would then just be a spurious phenomenon. To find out whether fluency per se can drive inferences about environmental variables, we experimentally manipulated fluency rather than measuring naturally existing fluency.

Method

Participants and design. Fifty students from the University of Basel participated in the study (36 women and 14 men; mean age = 24.1 years). Participants were presented with 49 pairs of cities. For each pair, participants were asked to infer which of the two cities has more residents (*inference task*). This task was identical to that in Study 3. Furthermore, each participant indicated which objects he or she recognized (*recognition task*). Students received a flat fee of CHF 15 (US\$13) or course credit.

Material. For experimentally induced fluency to have a reasonable chance of overcoming people's naturally existing retrieval fluency, we selected items that had similar degrees of pre-existing, relatively low levels of fluency. Specifically, we took advantage of Volz et al.'s (2006) pool of 525 cities drawn from different regions

⁷ Repeated measures analyses of variance showed linear trends in accordance rates over the four bins: cities environment, $F(1, 65) = 26.66$, $MSE = 0.079$, partial $\eta^2 = .29$; companies environment, $F(1, 61) = 24.15$, $MSE = 0.080$, partial $\eta^2 = .28$; and music artists, $F(1, 61) = 25.55$, $MSE = 0.066$, partial $\eta^2 = .30$.

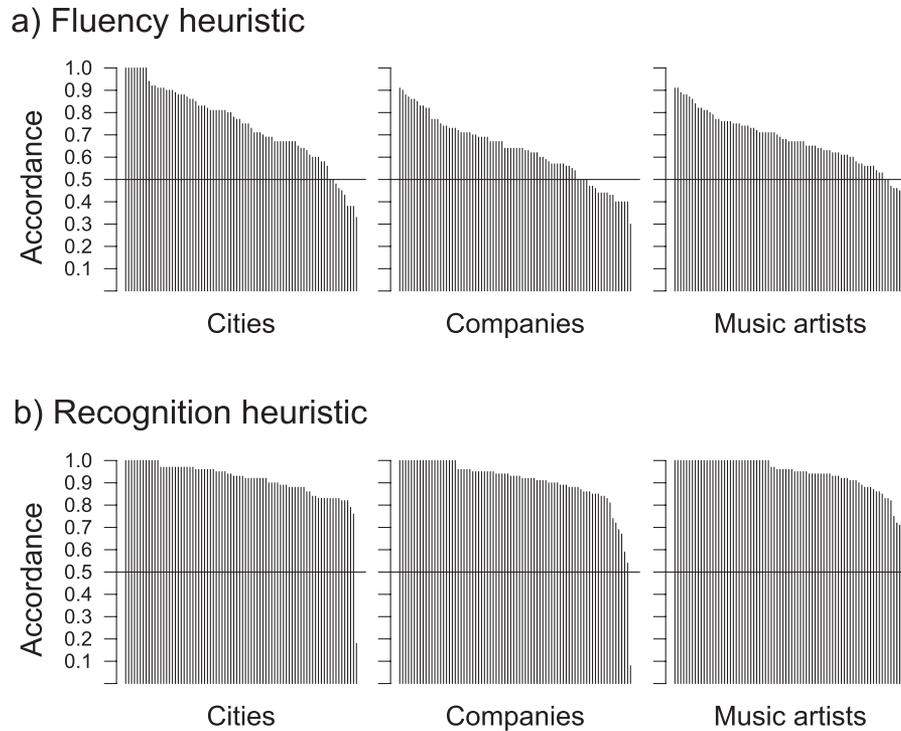


Figure 4. Accordance to the fluency and recognition heuristic. Percentage of inferences consistent with the fluency heuristic (Panel a) and recognition heuristic (Panel b) for participants in the cities, companies, and music artists environments, respectively. The individuals are ordered from left to right according to how often their judgments agreed with the respective heuristic. The horizontal line at 0.5 represents chance level accordance.

of the world. From this pool we first removed 10% of cities with the longest names (thus reducing uncontrolled variance in recognition time due to vastly different reading times). Then, we removed all cities that were not recognized by at least two thirds of those participants in a previous study, to increase the chance of obtaining pairs of cities in which both cities are recognized and the fluency heuristic is thus applicable. Finally, we winnowed down the pool further by splitting it into two halves according to recognition latency; the subset of the fastest recognized items was excluded to avoid ceiling effects in recognition latency. By using this procedure, we arrived at 68 cities (*target items*), with about equally long names and substantially reduced variance in pre-existing fluency (i.e., recognition latency). In addition, we randomly sampled from the set of excluded items 15 of the quickly recognized cities and 15 rarely recognized cities to form 15 filler pairs as well as 100 cities to serve as fillers in the fluency manipulation task (see below).

We assigned the 68 target items to two sets. They had comparable (a) recognition rates, (b) average median recognition speeds, (c) average number of letters in the names, and (d) average population sizes. Finally, we constructed pairs of cities by randomly selecting one city from each of the two sets of target items, thus arriving at 34 pairs. This whole procedure was implemented 25 times, thus creating 25 different lists of 34 pairs. Each of the 25 lists was administered in two versions. In one version, city *X* in each pair was selected for the fluency manipulation; in the other version, city *Y* was selected. That is, in each pair, one item was the

experimental, the other the control item. By thus counterbalancing the manipulated city within each pair, the fluency manipulation was not confounded with any idiosyncrasy of the constructed pairs.

Our fluency manipulation was a syllable counting task. Participants saw a sequence of 134 names of cities in a random order (the 34 target cities for which we intended to boost fluency, and the 100 filler city items) and were asked to judge whether each name had an even or uneven number of syllables. This fluency manipulation task was adapted from Yonelinas (2001, Experiment 3).

Procedure. After an introductory text explaining the syllable task, participants were presented with a sequence of 134 names of cities. By pressing one of two keys, people indicated whether the city name consisted of an even or uneven number of syllables. Following the syllable task, there was a 1-min break after which people worked on the inference task (as implemented in Study 3), involving 34 pairs of target items and 15 filler pairs. Once the inference task was completed, participants took the recognition task (as implemented in Study 3). The recognition judgments were used to identify those pairs in which a respondent recognized both items, a precondition for the fluency heuristic. Finally, participants' memories of the experimental manipulation were probed. Specifically, they indicated for each of the 68 target cities whether they thought that it was included in the syllable task at the outset of the study. To avoid invalid assumptions about the base rates, they were informed that 34 of the 68 cities were indeed included. After indicating whether a city was or was not included, they were asked to indicate on a 3-point scale how confident they were in

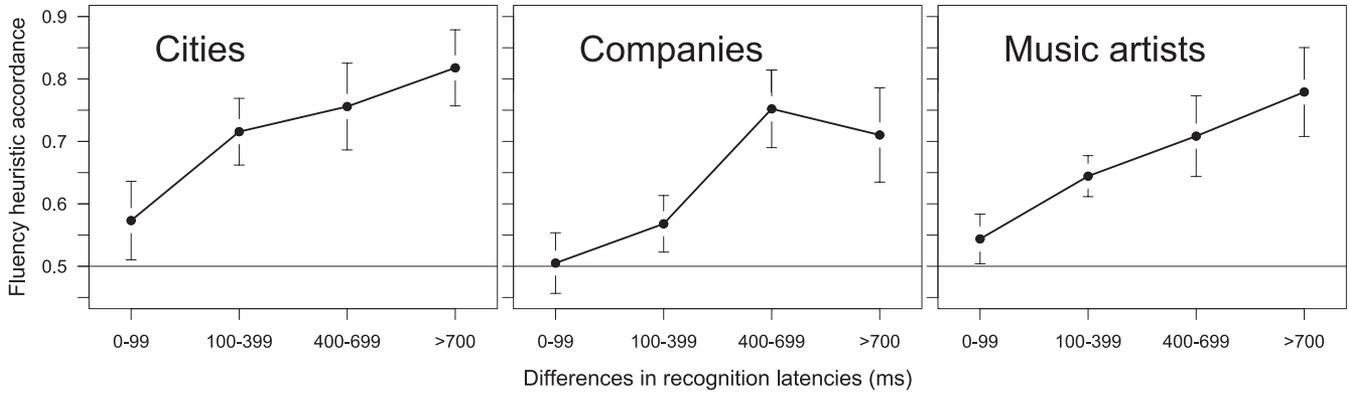


Figure 5. Increasing fluency heuristic accordance as a function of increasing differences in recognition latencies. Differences in objects *a* and *b*'s recognition latencies were classified into four bins. Then, we calculated the mean fluency accordance per bin (across participants). Error bars indicate 95% confidence intervals. The analyses reported in Footnote 7 show a significant linear trend across bins. The horizontal line at 0.5 represents chance level accordance.

their assessments. The rationale for monitoring memory of the experimental manipulation was that if people attributed their senses of fluency to the experiment rather than to their naturally acquired senses of fluency, they would be less likely to base their inferences on fluency (e.g., Jacoby & Whitehouse, 1989; Lombardi, Higgins, & Bargh, 1987).

Results

How did the fluency manipulation affect people's inferences of city size? To answer this question, we computed people's tendency to select the "manipulated" city (e.g., the city processed in the syllable task) over the non-manipulated one. Across all respondents, in .55 of all inferences, the manipulated city was chosen over the non-manipulated one (*SD* = .13; *CI* = .52, .59), one-sample *t* test against .5: *t*(49) = 2.94, *p* = .005. According to

Cohen's (1988) classification, this effect of the fluency manipulation amounts to a small to medium size (*d* = .37). In addition, double as many people had a fluency adherence rate above .5 than below .5 (28 vs. 14; binomial test: *p* = .02).

The effect of fluency is moderated by people's memory of the experimental manipulation. Based on each individual's 68 signal detection judgments and his or her confidence ratings (see above), we calculated *A_z*, fitting a parametric smooth receiver operating characteristics (ROC) curve (ROCFIT; see Metz, Shen, Wang, & Kronman, 1994; using JROCFIT; Eng, n.d.). *A_z* is a measure of how well people can discriminate between signal and noise (technically speaking, it is the area under the ROC curve). The average *A_z* value was .66 (*SD* = .08; *CI* = .64, .68.), suggesting that the ability to discriminate is better than chance (.5) but also far from perfect (1.0). Figure 7 plots people's fluency adherence as a function of *A_z*. The slope of the robust regression line suggests that those who were poorer at remembering which items had been part of the syllable task had a higher adherence rate than did those who remembered better (slope = -0.31; *SE* = 0.19, *p* = .06, one-tailed).

A final analysis concerns the combined effect of experimental and pre-experimental fluency. By using recognition speed data from a previous study, we determined that the faster recognized cities were chosen in Study 4 to be larger, on average, in 55% of cases. How did the experimentally induced fluency change this proportion? If the manipulated city was indeed the one recognized faster than the non-manipulated one (based on the previous results), fluency adherence grew from .55 to .60 (*SD* = .12; *CI* = .57, .63). In contrast, if the manipulated city was the one recognized slower than the non-manipulated one (based on the previous results), fluency adherence grew from .45 to .49 (*SD* = .13; *CI* = .45, .52).

To conclude, Study 4 shows that experimentally manipulated fluency shapes inferences. This result is compatible with the vast literature on the influence of fluency on preferences (see, e.g., Alter & Oppenheimer, 2007, for a review) and with previous results reported by, for instance, Kelley and Lindsay (1993), who demonstrated experimentally the impact of fluency of confidence

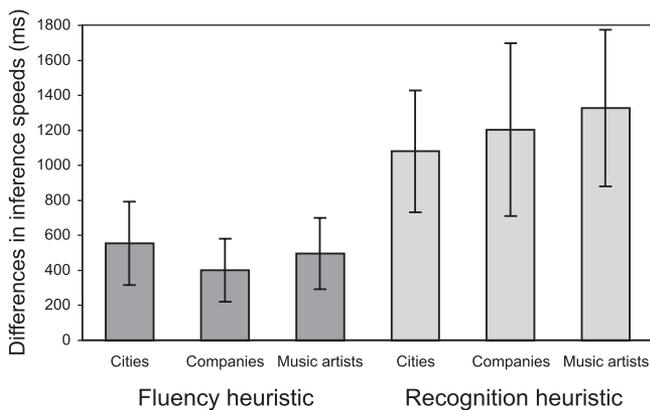


Figure 6. Inferences consistent with the fluency and recognition heuristics require less time. Bars represent the differences between median response times for inferences inconsistent and consistent with the fluency (left) and recognition heuristic (right), respectively. Positive differences indicate that inconsistent inferences took longer than consistent inferences. Error bars indicate 95% confidence intervals.

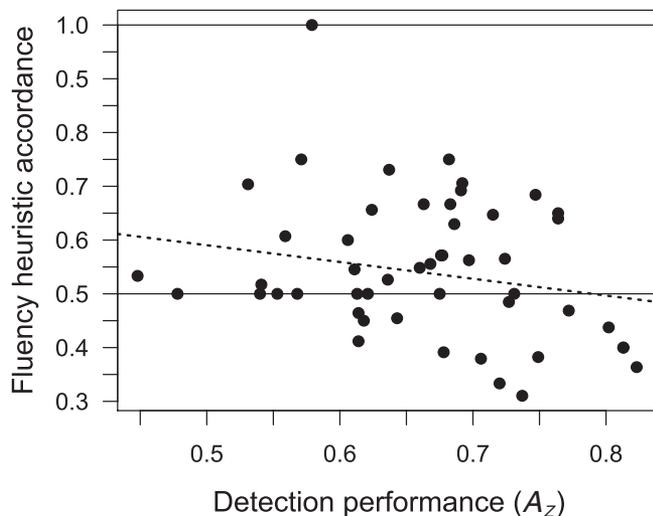


Figure 7. Fluency heuristic accordancy as a function of the ability to remember correctly which items were subjected to the experimental manipulation of fluency. A_z is the area under the ROC curve and can be interpreted as the probability of a correct response. The line represents the robust regression of the fluency heuristic accordancy as a function of the ability to correctly remember which items were included in the syllable task. The horizontal lines at 0.5 and 1.0 represent chance level and maximum accordancy, respectively.

in the domain of general knowledge questions. The effects in Study 4 are small to medium in size, but it is worth keeping in mind that our experimental manipulation of fluency consisted of merely the judgment of whether or not a city name has an even or uneven number of syllables. Thus it was just the proverbial drop in the ocean compared to the lifetime of exposure to these city names.

General Discussion

Automatically generated in the process of retrieval from memory, recognition and retrieval fluency are two pieces of information that are likely to precede the arrival of many other probabilistic cues retrieved from memory. The recognition heuristic and the fluency heuristic are two mind tools that have been proposed to exploit both pieces of mnemonic information. In several studies, we have investigated the fluency heuristic side by side with the recognition heuristic. In Study 1, we found that retrieval fluency can be a valid predictor of objective properties of the world, but the degree to which it is indicative of the criterion varies across environments. To the best of our knowledge, this is one of the first systematic demonstrations of the ecological validity of fluency information, based on the chain of correlations between criterion, mediators, and retrieval fluency within precisely defined reference classes. In Study 2, we examined whether people can reliably discriminate differences in retrieval fluency, a prerequisite for employing the fluency heuristic. Indeed, latencies that differ by more than 100 ms can be distinguished. The larger the differences, the better people can discriminate and the larger fluency validity proved to be. In Study 3, we found that between about three fourths and two thirds of the inferences, in which the fluency heuristic could be applied, conformed to the heuristic. Further-

more, accordance with the fluency heuristic increased with larger differences in recognition latencies—a manifestation of ecological rationality insofar as retrieval fluency is more likely to yield accurate inferences with larger differences (given an ecological correlation). Finally, in Study 4, we experimentally manipulated fluency and found that it had a direct impact on inferences.

In what follows, we turn to a possible explanation of the robust observation that recognition validity surpasses fluency validity, and we discuss what people can gain from the fluency heuristic and under what circumstances it may be used.

Why is Recognition More Valid Than Fluency Information?

In our studies, recognition validity consistently exceeded fluency validity. Why is that? Recognition and retrieval fluency both depend on activation, which encodes environmental frequencies, which, in turn, can be associated with the criterion (ecological validity). Within ACT-R, the difference between recognized and unrecognized objects maps onto the difference between items that exceed a retrieval threshold and those that fall below. For a moment, let us treat the threshold as the median value of a predictor variable that splits the criterion distribution into two parts. On average, recognized objects will score higher on the criterion than will unrecognized ones, just as items above the median will, on average, have a higher criterion value than those below. The fluency heuristic is akin to restricting the calculation of a correlation between predictor and criterion to those items that score above the median of the predictor variable. In all likelihood, the correlation and, by extension, the fluency validity will be lower relative to calculating the correlation across the whole range of the criterion value. Because the fluency heuristic requires both objects to be recognized, the relevant range of objects' activations is restricted, and, consequently, the link between activation and criterion is less pronounced. In contrast, the recognition heuristic requires one object to be recognized and the other not. Consequently, the range of activations exceeds that for the fluency heuristic, and therefore the link between activation and criterion will be more pronounced, relative to the fluency heuristic. On this view, recognition validities exceed fluency validities, because the fluency heuristic, unlike the recognition heuristic, applies to objects whose activations exceed the retrieval threshold. Metaphorically speaking, the fluency heuristic deals with a world full of shades of white, whereas the recognition heuristic faces a black-and-white world.

What Do People Gain From Relying On Fluency?

When this question is interpreted simply with respect to accuracy, the answer depends on which of two benchmarks is used. If the alternative to the fluency heuristic is guessing (for example, in cases in which no further knowledge is available), the heuristic clearly tops chance performance. As Study 1 demonstrated, fluency validity surpasses chance performance, ranging from .58 to .66 (Figure 1). Although one would hesitate to go mushroom picking with such accuracy, one could make a comfortable living picking stocks. If the alternative to relying on fluency is mustering additional cue knowledge, however, the performance of the fluency heuristic depends on the accuracy of other decision strategies

that depend on this knowledge (see Gigerenzer & Goldstein, 1996, for candidate strategies). Should the validity of these knowledge-based strategies exceed that of fluency, by using fluency a person sacrifices some accuracy.

We have no direct measure of the validity of these knowledge-based strategies but we can use a proxy to extrapolate it, namely, the level of accuracy in those cases in which inferences are inconsistent with the fluency heuristic. In all three environments in Study 3, the average fluency validity was as high or higher than the average accuracy of additional knowledge: .66 vs. .59 ($Mdn = .67$, $SD = .34$), .58 vs. .47 ($Mdn = .5$, $SD = .26$), and .58 vs. .52 ($Mdn = .5$, $SD = .25$) in the cities, companies, and music artists environments, respectively. These numbers stem from a selective set of items and thus need to be interpreted with care. Yet, they provide a first indication that even when compared with the accuracy of knowledge-based strategies the fluency heuristic stands its ground in the environments investigated. At the same time, let us emphasize that there are domains in which reliance on fluency can undoubtedly result in disadvantageous decisions. Companies, for instance, pay great sums to influence the public's recognition and fluency of product names (Goldstein, 2007). No doubt, the more fluent product name need not always signal the better product.

In any consideration of what the fluency heuristic has to offer, two of its properties deserve attention. First, the heuristic is least likely to be used in cases in which its validity hits bottom, that is, when people find it difficult to tell the recognition latencies apart (see Figures 2 and 3). In the cities environment, for instance, the adherence rate in the 50% of inferences with the smallest differences in recognition latencies is .62, relative to .77 in the 50% of inferences with the largest differences. The corresponding fluency validities for the smaller and larger differences are .56 and .69, respectively. In other words, the heuristic's requirement that the recognition latencies be discernible inures the user from using fluency when it is least beneficial. Second, the wisdom of the fluency heuristic should be evaluated in light of the fact that the retrieval of other cue knowledge is effortful and time consuming. In other words, the decision maker faces an effort-accuracy tradeoff (Payne, Bettman, & Johnson, 1993). Using the fluency heuristic may not be the best people can do, but it enables them to arrive at inferences swiftly and to surpass chance performance if they have no other cue knowledge; in addition, they are protected from using it when it is least appropriate.

When Do People Resort to the Fluency Heuristic?

In light of the fluency heuristic's competitors (Gigerenzer & Goldstein, 1996), what are the circumstances under which people use it? Some critics of the adaptive toolbox metaphor suggest that the issue of strategy selection runs the risk of making the dubious assumption of some kind of omniscient homunculus or über-heuristic that selects from the toolbox (e.g., Newell, 2005). However, possible solutions to the thorny issue of strategy selection have been proposed that make do without a mysterious homunculus. One solution is that people learn to match specific heuristics to specific statistical structures in the world through feedback (Rieskamp & Otto, 2006). Alternatively, the various mind tools in the toolbox may be arranged according to the same criteria that the ACT-R framework (Anderson et al., 2004; Anderson & Lebiere,

1998) uses in order to determine the activation strength of a memory record, namely, recency and frequency of occurrence. That is, the various tools may be ordered such that the most frequently and most recently used tool is examined first; if it does not enable an inference, the next tool in the hierarchy will be examined, and so on.

There is still another, albeit not exclusive, route to strategy selection. Knowledge and time, or lack thereof, may also guide the selection process. Recently, Marewski and Schooler (2007) found that the fluency heuristic appears most likely to be used when both objects are recognized and no other probabilistic cue knowledge is available. When knowledge was available, a knowledge-based strategy described people's inferences better than the fluency heuristic. This result suggests the possibility that people may be inclined (a) to use the recognition heuristic if one object is recognized and the other is not (see Figure 4b) and if people have no definite and conclusive knowledge of the target variable that renders use of cues unnecessary (Pachur & Hertwig, 2006); (b) to use the fluency heuristic if both objects are recognized and no other knowledge (e.g., in terms of probabilistic cues) is available; and (c) to use knowledge-based strategies when both objects are recognized and additional knowledge (e.g., in terms of probabilistic cues) is available.

Let us end with a clarification. We have specified the fluency heuristic in terms of a condition-action production ("If-then") rule, so that it can be instantiated as a computer program. This specification, however, does not mean that we have already completely understood which antecedence conditions need to be met for the action to be executed. Our studies have confirmed that the difference between recognition latencies needs to be at least 100 ms, otherwise the action cannot not be executed (e.g., Figures 3 and 5). Future investigations will help to clarify which additional conditions need to be met. Such other conditions may relate, for instance, to the presence or absence of other probabilistic cue knowledge, or to the validity of present cue knowledge.

Conclusion

David Hume (1740/1978) believed that the mind meticulously tallies frequencies and apportions degrees of belief accordingly. Today, we know that Hume endowed the human mind with too exquisite a sensitivity to frequencies. Yet, as various memory models assume, there is a mapping between environmental frequencies and mnemonic information such as activation strengths. The fluency heuristic is one mind tool that artfully and swiftly probes memory for the encapsulated frequency information that can veridically reflect statistical regularities in the world.

References

- Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 9369-9372.
- Alter, A. L., & Oppenheimer, D. M. (2007). *Uniting the tribes of fluency into a metacognitive nation*. Manuscript submitted for publication.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111, 1036-1060.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

- Benjamin, A. S., Bjork, R. A., & Hirshman, E. (1998). Predicting the future and reconstructing the past: A Bayesian characterization of the utility of subjective fluency. *Acta Psychologica*, *98*, 267–290.
- Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*, 55–68.
- The best-paid athletes. (2004, June 24). Retrieved July 28, 2004, from the *Forbes* Web site: <http://www.forbes.com/2004/06/23/04athletesland.html>
- Bröder, A., & Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica*, *121*, 275–284.
- Butler, R. (2003). *Cities and urban areas in United States with population over 100,000*. Retrieved July 10, 2008, from http://www.mongobay.com/cities_urban_01.htm
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Erlbaum.
- Die 100 grössten Unternehmen Deutschlands [The 100 biggest German companies]. (n.d.). Retrieved August 10, 2004, from <http://www.sueddeutsche.de/imperia/md/content/pdf/wirtschaft/tabelle.pdf>
- Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). Minerva-DM: A memory processes model for judgments of likelihood. *Psychological Review*, *106*, 180–209.
- Eng, J. (n.d.). *ROC analysis: Web-based calculator for ROC curves*. Retrieved June 22, 2004, from <http://www.jrocfit.org>
- Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, *35*, 1–36.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Goldstein, D. G. (2007). Getting attention for unrecognized brands. *Harvard Business Review*, *85*, 24–28.
- Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75–90.
- Hertwig, R., Pachur, T., & Kurzenhäuser, S. (2005). Judgments of risk frequencies: Tests of possible cognitive mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 621–642.
- Hume, D. (1978). *A treatise of human nature*. Oxford, England: Clarendon Press. (Original work published 1740).
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306–340.
- Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, *118*, 126–135.
- Kelley, C. M., & Jacoby, L. L. (1998). Subjective reports and process dissociation: Fluency, knowing, and feeling. *Acta Psychologica*, *98*, 127–140.
- Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*, 1–24.
- Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, *52*, 478–492.
- Lombardi, W. J., Higgins, E. T., & Bargh, J. A. (1987). The role of consciousness in priming effects on categorization: Assimilation versus contrast as a function of awareness of the priming task. *Personality and Social Psychology Bulletin*, *13*, 411–429.
- Marewski, J. N., & Schooler, L. J. (2007, August). *The cognitive niche of the fluency heuristic*. Paper presented at the 21st Biennial Conference on Subjective Probability, Utility and Decision Making, Warsaw, Poland.
- Metz, C. E., Shen, J. H., Wang, P. L., & Kronman, H. B. (1994). *ROCFIT software*. Technical Report, University of Chicago, Chicago, IL.
- Newell, B. R. (2005). Re-visions of rationality? *Trends in Cognitive Sciences*, *9*, 11–15.
- Newell, B. R., & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, *19*, 333–346.
- Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition*, *90*, B1–B9.
- Pachur, T., & Biele, G. (2007). Forecasting from ignorance: The use and usefulness of recognition in lay predictions of sports events. *Acta Psychologica*, *125*, 99–116.
- Pachur, T., Bröder, A., & Marewski, J. N. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making*, *21*, 183–210.
- Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 983–1002.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, England: Cambridge University Press.
- Pohl, R. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, *19*, 251–271.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*, 364–382.
- Reimer, T., & Katsikopoulos, K. V. (2004). The use of recognition in group decision-making. *Cognitive Science*, *28*, 1009–1029.
- Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 150–162.
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207–236.
- Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, *112*, 610–628.
- Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, *14*, 332–348.
- Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, *61*, 195–202.
- Schwarz, N., & Wänke, M. (2002). Experiential and contextual heuristics in frequency judgments: Ease of recall and response scales. In P. Sedlmeier & T. Betsch (Eds.), *Etc.: Frequency processing and cognition* (pp. 89–108). New York: Oxford University Press.
- Sedlmeier, P., Hertwig, R., & Gigerenzer, G. (1998). Are judgments of the positional frequencies of letters systematically biased due to availability? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 754–770.
- Serwe, S., & Frings, C. (2006). Who will win Wimbledon? The recognition heuristic in predicting sports events. *Journal of Behavioral Decision Making*, *19*, 321–332.
- Shaw, J. S., III. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied*, *2*, 126–146.
- Shaw, J. S., III, McClure, K. A., & Wilkens, C. E. (2001). Recognition instructions and recognition practice can alter the confidence–response time relationship. *Journal of Applied Psychology*, *86*, 93–103.
- Top artists. (2003). Retrieved July 10, 2008, from <http://web.archive.org/web/20040216195220/http://www.ria.com/gp/bestsellers/topartists.asp>
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

- Tversky, A., & Kahneman, D. (1974, September 27). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.
- Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 219–230.
- Volz, K. G., Schooler, L. J., Schubotz, R. I., Raab, M., Gigerenzer, G., & von Cramon, D. Y. (2006). Why you think Milan is larger than Modena: Neural correlates of the recognition heuristic. *Journal of Cognitive Neuroscience*, *18*, 1924–1936.
- Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1235–1253.
- Whittlesea, B. W. A., & Leboe, J. P. (2003). Two fluency heuristics (and how to tell them apart). *Journal of Memory and Language*, *49*, 62–79.
- Winkielman, P., Schwarz, N., Fazendeiro, T. A., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Mahwah, NJ: Erlbaum.
- The world's richest people. (2004). Retrieved July 29, 2004, from the *Forbes* Web site: <http://www.forbes.com/2004/02/25/bill04land.html>
- Yonelinas, A. P. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General*, *130*, 361–379.

Received February 9, 2007

Revision received April 25, 2008

Accepted May 5, 2008 ■

Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write to the address below. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, “social psychology” is not sufficient—you would need to specify “social cognition” or “attitude change” as well.
- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

Write to Journals Office, American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242.